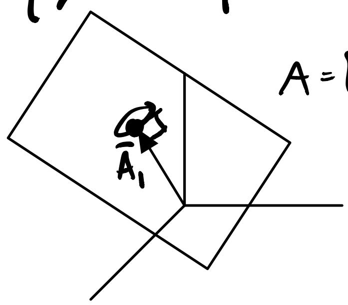
AFFINE SPACES:

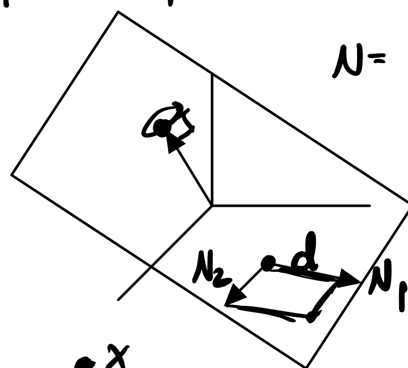
2 representations

$$\{x \in \mathbb{R}^n \mid Ax = b\} = \{x \in \mathbb{R}^n \mid x = Nz + d\}$$



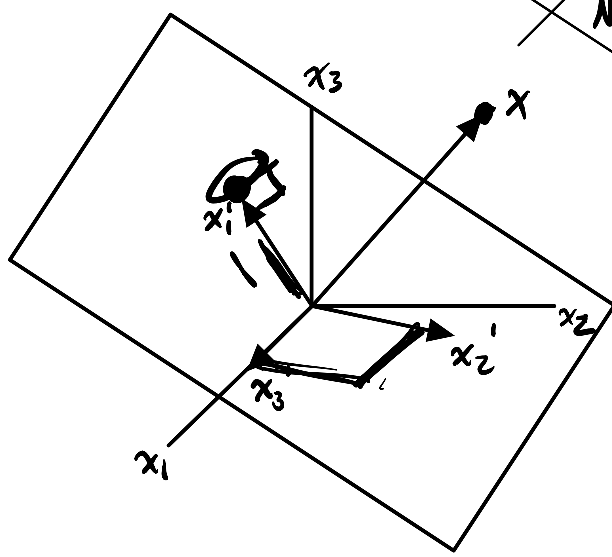
$A = [\ \bar{A}_1^T\ ]$

normal direction



$N = [N_1\ N_2]$

"basis" for affine space



$$x = [A^T N]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$x = [\bar{A}_1\ N_1\ N_2]\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}\begin{matrix} \} z_1 \\ \} z_2 \end{matrix}$$

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax = b$$

$$\frac{\partial f}{\partial x} = x^T Q = -v^T A \quad \left(\text{for } f(x) = \tfrac{1}{2}x^T Q x\right)$$

$$Ax = b$$

solve for x & v

$$\min_{\substack{x \in \mathbb{R}^n \\ z \in \mathbb{R}^p}} f(x) \quad \text{s.t.} \quad x = Nz + d$$

$$N \in \mathbb{R}^{n \times p}$$

$$\min_{\substack{x\in\mathbb{R}^n \\ z\in\mathbb{R}^p}} f(x) \leftarrow$$

s.t. $\underline{x} = N\overset{\uparrow}{z} + d$

z unconstrained

$$\min_z f(Nz+d) \Rightarrow \frac{\partial f}{\partial z} = 0 \,,\ x = Nz+d$$

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial z} = \frac{\partial f}{\partial x} N = 0 \,,\ x = Nz+d$$

$$\min_{x,z} \frac{1}{2}x^TQx + c^Tx$$

s.t. $x = Nz+d$] → either treat this as a constraint or plug it in

Plugging in...

$$\frac{1}{2}(d^T+z^TN^T)Q(Nz+d) + c^T(Nz+d) = f(z)$$

$$\frac{1}{2}d^TQd + d^TQNz + \frac{1}{2}z^TN^TQNz + c^TNz + c^Td$$

$$\frac{\partial f}{\partial z} = z^TN^TQN + d^TQN + c^TN = 0$$

$$\boxed{z^T = (-d^TQN - c^TN)(N^TQN)^{-1}}$$

solve for x.... $x = Nz+d \leftarrow$

$$x = -N(N^TQN)^{-1}(N^TQd + N^Tc) + \underline{d}$$

treating as a constraint

$$\min_{x,z} \ \frac{1}{2}x^TQx + c^Tx$$

s.t. $x = Nz + d \Rightarrow [I, -N]\begin{bmatrix} x \\ z \end{bmatrix} = d$

$Ax = b$

$\begin{bmatrix} x \\ z \end{bmatrix}$ v^TA

x, z "dual variable" τ

or "lagrange multiplier"

$$\left.\begin{array}{l} \frac{\partial f}{\partial x} = \cancel{x^TQ + c^T = \tau^T[I, -N]} \\ \quad x = Nz + d \end{array}\right] \to \text{solve for } x, z \ \& \ \tau$$

$$\frac{\partial f}{\partial \begin{bmatrix} x \\ z \end{bmatrix}} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial z} \end{bmatrix} = \tau^T[I, -N]$$

$$\begin{bmatrix} x^TQ + c^T & 0 \end{bmatrix} = \tau^T[I, -N]$$

$$x = Nz + d$$

$x^TQ + c^T = \tau^T$, $0 = -\tau^TN$

τ needs to be $\perp$ to cols of N...

$AN = 0 \Rightarrow \tau = A^Tv$ for some v

$\Rightarrow \tau^T = v^TA$

$Ax = b$

$$\boxed{x^TQ + c^T = v^TA}$$

$x^TQ + c^T = \tau^T$

$\left[(d^T + z^TN^T)Q + c^T = \tau^T\right] \times N$

$$d^T Q N + z^T N^T Q N + c^T N = \tau^T \cancelto{0}{N}$$

$$\boxed{z^T = \left(-d^T Q N - c^T N\right)\left(N^T Q N\right)^{-1}}$$

Summery

- can treat constraints like $x = Nz + d$ either as constraints or plug them into $f(x)$
- if you treat them as constraints you get an extra lagrange multiplier τ but same solution for $z, x, \ldots$
- if you have two different formulations of the same constraint

$$Ax = b \iff x = Nz + d$$

Lagrange multi. v $\qquad$ τ

some relationship

$$v^T A = \tau^T$$

$$\frac{\partial f}{\partial x} = v^T A \qquad v \in \mathbb{R}^m$$

$$\left[\frac{\partial f}{\partial x} \;\; \frac{\partial f}{\partial z}\right] = \tau^T \left[I \; -N\right] \qquad \tau \in \mathbb{R}^n$$

$$\underline{\frac{\partial f}{\partial x} = \tau^T} \qquad \frac{\partial f}{\partial z} = \tau^T N$$

$$Ax = b \qquad 0 = \tau^T N$$

$A \in \mathbb{R}^{m \times n}$ $\quad m\left[\quad \overset{n}{\longleftrightarrow} \quad\right]$

τ

$\boxed{\frac{\partial f}{\partial x} = v^T A}$ $\qquad$ $\boxed{\frac{\partial f}{\partial x} = \tau^T}$ $\qquad$ $\boxed{0 = \tau^T N}$

$\tau \perp$ cols of N

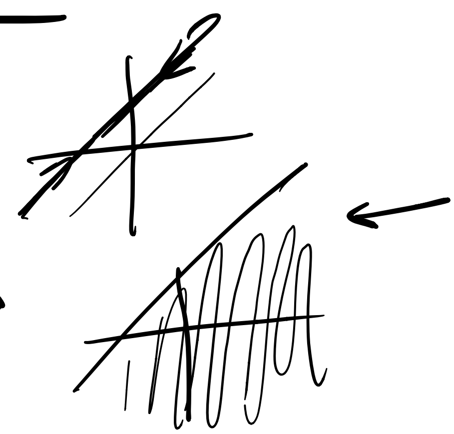
$\Downarrow$

$\frac{\partial f}{\partial x} \in R(A^T)$ $\qquad$ $\tau = A^T v$

or a lin comb of rows of A $\qquad$ $\frac{\partial f}{\partial x} = \tau \in R(A^T)$

INEQUALITY CONSTRAINTS:

$$\min_x f(x)$$
$$\text{s.t. } g(x) = 0$$
$$h(x) \geq 0$$



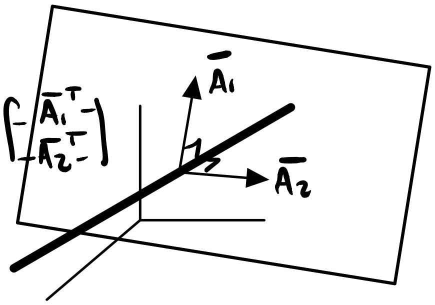
$\Rightarrow Ax = b \Leftarrow$

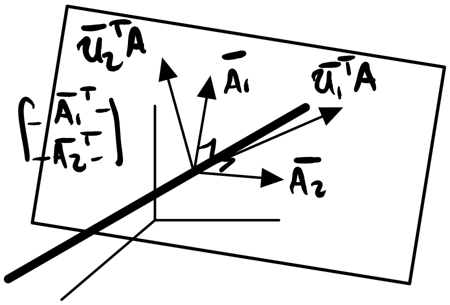
$\rightarrow \mathcal{U}$, invertible

$$\{x \mid Ax = b\} = \{x \mid \mathcal{U}Ax = \mathcal{U}b\}$$

rows of $\mathcal{U}A$ are lin combs of rows of A

$A = \begin{bmatrix} -\bar{A}_1^T- \\ -\bar{A}_2^T- \end{bmatrix}$





$uA = \begin{bmatrix} \bar{u}_1^T \\ \bar{u}_2^T \end{bmatrix} \begin{bmatrix} A \end{bmatrix}$

if $Cx \ge d$

$$\{x \mid Cx \ge d\} \ne \{x \mid uCx \ge ud\}$$

$x \in \mathbb{R}$
$C \in \mathbb{R}$
$d \in \mathbb{R}$
$u \in \mathbb{R}_+$

$\rightarrow$ $Cx \ge d \Rightarrow ucx \ge ud$

if $u < 0 \Rightarrow uCx \le ud$

if $u_{ij} \ge 0$ every element of u is ≥ 0

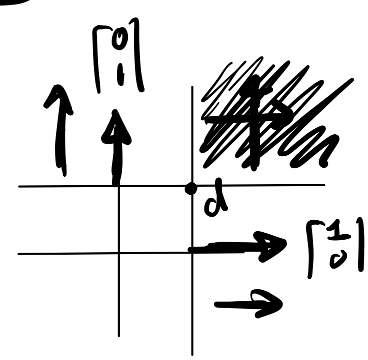
$\rightarrow$ if $\boxed{Cx \ge d} \Rightarrow \boxed{uCx \ge ud}$ $\leftarrow$

$\not\Leftarrow$

Ex. $x \in \mathbb{R}^2$

$$x \ge d$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \ge \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

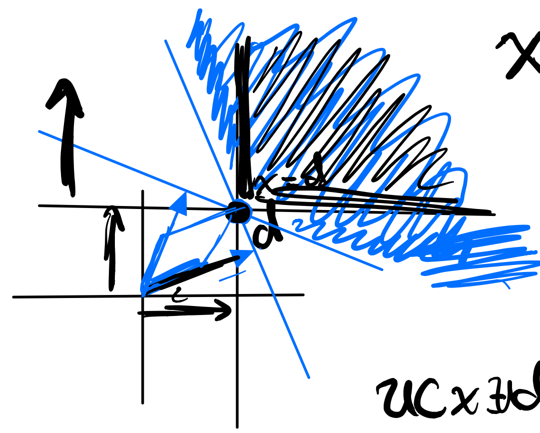


$\mathcal{U}x \geq \mathcal{U}d$

$\begin{bmatrix} -\bar{u}_1^T- \\ -\bar{u}_2^T- \end{bmatrix} x \geq d$

$\bar{u}_1^T x \geq d_1$

$\bar{u}_2^T x \geq d_2$



$\mathcal{U}Cx \not\geq \mathcal{U}d$

$\mathcal{U} \quad x \geq d$

$\mathcal{U}x = \boxed{\mathcal{U}d}$

$[\mathcal{U}x = \mathcal{U}d] \Rightarrow$

$x = d \rightarrow \mathcal{U}d = \mathcal{U}d$

$\mathcal{U}Cx = \mathcal{U}d$

$\mathcal{U}d \neq d$

d_1 in the direction $\begin{bmatrix}1\\0\end{bmatrix}$

d_2 in the direction $\begin{bmatrix}0\\1\end{bmatrix}$

$\begin{bmatrix}d_1'\\d_2'\end{bmatrix} = \mathcal{U}d$

d_1' in the direction $\bar{u}_1$

d_2' in the direction $\bar{u}_2$

$\mathcal{U}_{ij} \geq 0 \not\Rightarrow \mathcal{U}^{-1}_{ij} \geq 0$

$d = \mathcal{U}^{-1} z \qquad x = \mathcal{U}^{-1}$

Suggestion: play with this...

desmos.com

geogebra.org

Slack Variables

$\rightarrow \quad Cx \geq d \quad \Longrightarrow \quad Cx = d + s$

$Cx - s = d \quad , \quad \underline{s \geq 0}$

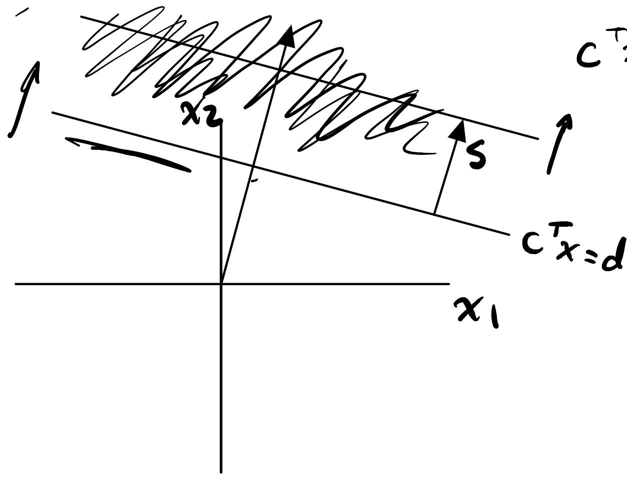
inequality const. $\qquad$ equality const. $\qquad$ inequality

adds a slack variable s.

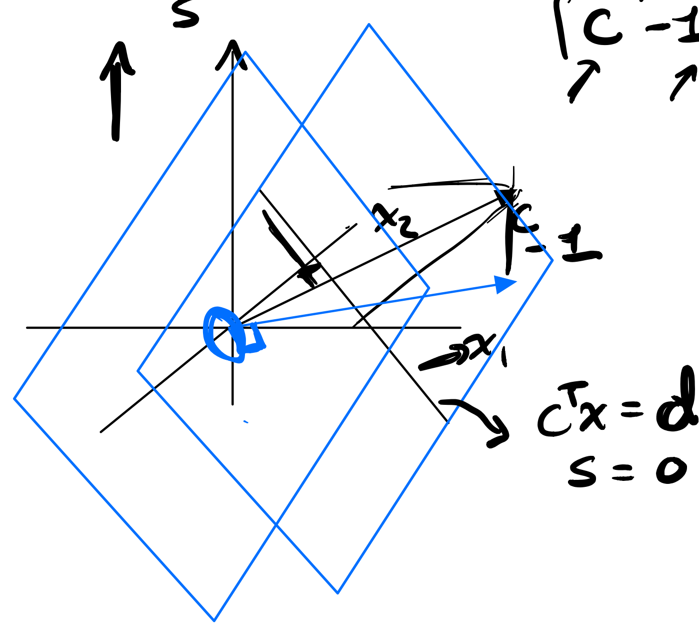
Ex. $c^T x \geq d$

$c^T x - s = d \Rightarrow c^T x - d = s$

$c^T x - d = s$



$$\begin{bmatrix} c^T & -1 \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = d \quad s \geq 0$$



$$Cx - s = d$$

for $C \in \mathbb{R}^{m \times n}$ $x \in \mathbb{R}^{n}$, $\underline{s \in \mathbb{R}^{m}}$

one slack variable for ea. constraint

when $\underset{\sim}{A}x = b$

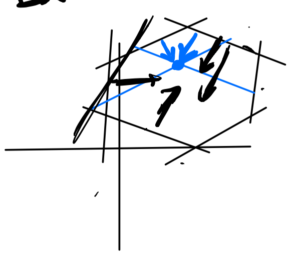
always fat

$Cx = d$

$Cx \geq d$

can be tall or fat

Ex. C tall $\Rightarrow C \in \mathbb{R}^{6 \times 2}$



slack variables

$Cx - s \geq d$

$$\begin{array}{cc} & \begin{array}{cc} 2 & 6 \end{array} \\ 6 & \underbrace{\begin{bmatrix} C & -I \end{bmatrix}}_{\substack{\text{always fat} \\ 6 \times 8}} \end{array} \begin{bmatrix} x \\ s \end{bmatrix} = d \qquad \underline{s \geq 0}$$

$Cx \geq d$

allowed to multiply by a positive diagonal matrix

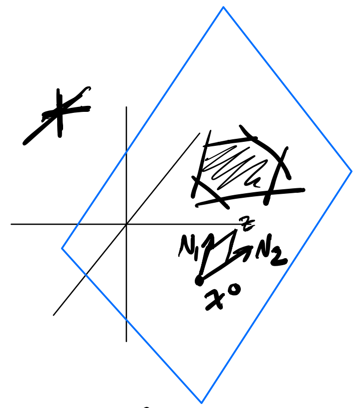
$$C = \begin{bmatrix} \bar{c}_1^T \\ \vdots \\ \bar{c}_m^T \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{|\bar{c}_1|} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{|\bar{c}_m|} \end{bmatrix} Cx \geq d$$

$$\rightarrow \begin{bmatrix} \bar{c}_1^T / |\bar{c}_1| \\ \bar{c}_m^T / |\bar{c}_m| \end{bmatrix} x = \begin{bmatrix} d_1 / |\bar{c}_1| \\ \vdots \\ d_m / |\bar{c}_m| \end{bmatrix} + \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix}$$

$\left[Ax = b \quad Cx \geq d \right]$ ✗

fat A

$Ax = b \Rightarrow x = Nz + x^0$

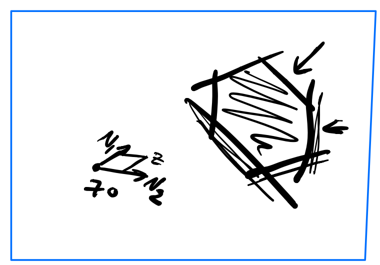


$x = Nz + x^0$ z unconstrained

$Cx \geq d$

$CNz + Cx^0 \geq d$

$-CNz \geq d - Cx^0$



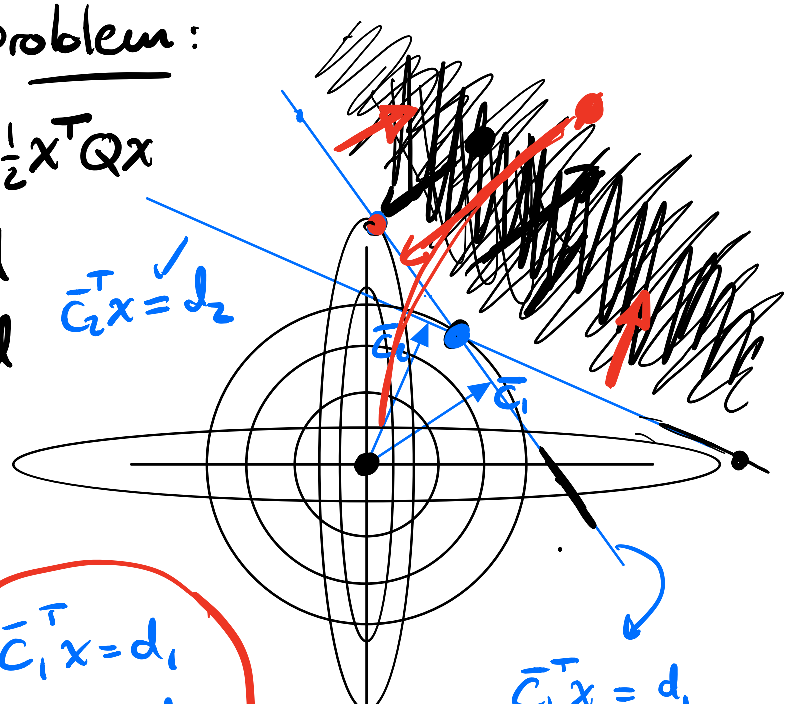
optimization problem:

$$\min_x \; f(x) = \tfrac{1}{2} x^T Q x$$

$$\text{s.t.} \quad Cx \geq d$$

$$\left[\vdots \;\; \vdots \right] x = d$$

$\bar{c}_2^T x = d_2$ ✓



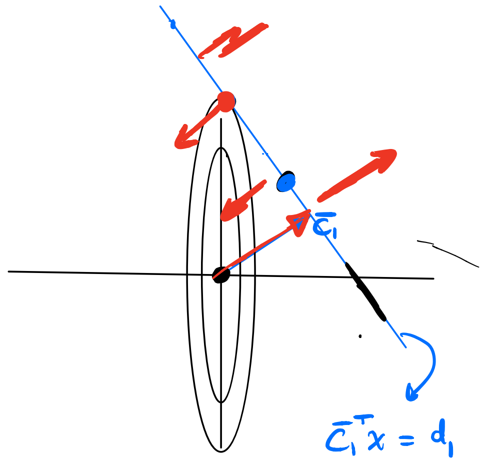
$C = \begin{bmatrix} \bar{c}_1^T \\ \bar{c}_2^T \end{bmatrix}$

$d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$

$\bar{c}_1^T x = d_1$

$\bar{c}_2^T x > d_2$

$\bar{c}_1^T x = d_1$



Before:

optimality conditions:

$$\frac{\partial f}{\partial x} = v^T A \qquad Ax = b$$

$$\downarrow \qquad\qquad \downarrow$$

$$\left[\frac{\partial f}{\partial x} = \underline{\mu}^T C \qquad Cx \geq d \right]$$

doesn't work

need new optimality conditions

$\mu \in \mathbb{R}^m$

$$\frac{\partial f}{\partial x} = \mu^T C$$

$\mu \geq 0 \longrightarrow$ the constraints can only push in one direction $\rightarrow \mu_i \geq 0$

$Cx \geq d$

$\rightarrow \mu_i [\underbrace{Cx - d}_{s_i}]_i = 0 \longrightarrow$ acts as a switch to turn the push back of the constraints on and off.

with slack variables...

$$\min_{x,s} f(x)$$

s.t. $Cx - s = d$, $s \geq 0$

$$\begin{array}{l} \tau \rightarrow \\ \rightarrow \mu \rightarrow \end{array} \begin{bmatrix} C & -I \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} \begin{array}{c} = d \\ \geq 0 \end{array}$$

$C \in \mathbb{R}^{m \times n}$ $\quad \tau \in \mathbb{R}^m$ $\quad \mu \in \mathbb{R}^m$

$$\begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial s}\end{bmatrix} = \begin{bmatrix}\tau^T & \mu^T\end{bmatrix}\begin{bmatrix} C & -I \\ 0 & I\end{bmatrix} = \tau^T[C, -I] + \mu^T[0 \ I]$$

$Cx - s = d$, $s \geq 0$ $\qquad \mu_i \underline{s_i} = 0$

no constraints on τ $\qquad \mu \geq 0$

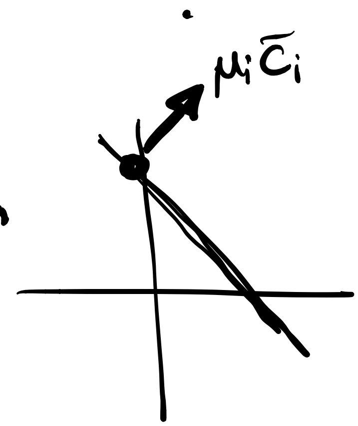
$\mu_i \geq 0$

$s_i \geq 0$ $\qquad \mu_i s_i = 0$

if $s_i = 0 \Rightarrow \mu_i s_i = 0$

so μ_i can be > 0

if $s_i > 0 \Rightarrow \mu_i = 0$



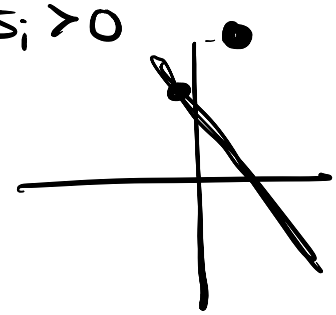
μ_i is the push back of constraint i

if constraint i has no slack then μ_i can push back against $\frac{\partial f}{\partial x}$

but $\mu_i \geq 0$ (pushing in only one direction)

if constraint i has slack... ie. $s_i > 0$

then μ_i can't push back



$\mu_i = 0$

$\mu_i s_i = 0$ $\qquad$ called a complementary slackness constraint

$$\mu_i s_i = 0 \quad \forall i \iff \mu^T s = \sum_i \mu_i s_i = 0$$

$$\mu_i \geq 0 \quad \forall i \qquad\qquad \mu \geq 0$$

$$s_i \geq 0 \quad \forall i \qquad\qquad s \geq 0$$

$$\min_{x,s} \ f(x) \qquad A \in \mathbb{R}^{m \times n} \quad C \in \mathbb{R}^{p \times q}$$

$$\text{s.t.} \quad Ax = b, \quad Cx - s = d \qquad s \geq 0$$

$$\underbrace{\begin{bmatrix} A & 0 \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = b}_{v \in \mathbb{R}^m} \quad \underbrace{\begin{bmatrix} C & -I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = d}_{\tau \in \mathbb{R}^p} \quad \underbrace{\begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} \geq 0}_{\mu \in \mathbb{R}^p}$$

for unconstrained $\frac{\partial f}{\partial x} = 0$

for constrained...

$$\star \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial s} \end{bmatrix} = \underline{v}^T \begin{bmatrix} A & 0 \end{bmatrix} + \underline{\tau}^T \begin{bmatrix} C & -I \end{bmatrix} + \underline{\mu}^T \begin{bmatrix} 0 & I \end{bmatrix}$$

$$\rightarrow \mu \geq 0 \qquad \underline{\mu_i s_i} = 0 \iff \mu^T s = 0 \leftarrow$$

$$Ax = b, \quad Cx - s = d, \quad s \geq 0 \qquad \star$$

solve for x, s, v, τ, μ

KKT CONDITIONS $\leftarrow$ [optimality conditions

$\star \quad \frac{\partial f}{\partial x} = v^T A + \tau^T C \qquad \frac{\partial f}{\partial s} = 0 = -\tau^T = \underline{\mu^T}$

Generally

$$\frac{\partial f}{\partial x} = 0$$

$$\min_x \quad f(x)$$

$$\text{s.t.} \quad g(x) = 0 \} \rightarrow v \quad \leftarrow \quad \frac{\partial \mathcal{L}}{\partial (x, v, \mu)} = 0$$

$$h(x) \geq 0 \} \rightarrow \mu$$

Lagrangian:

$$\mathcal{L}(x, v, \mu) = f(x) + v^T g(x) - \mu^T h(x) \leftarrow$$

KARUSH-KHUN-TUCKER (KKT) CONDITIONS

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} + v^T \frac{\partial g}{\partial x} - \mu^T \frac{\partial h}{\partial x} = 0$$

$\leftarrow$ stationarity

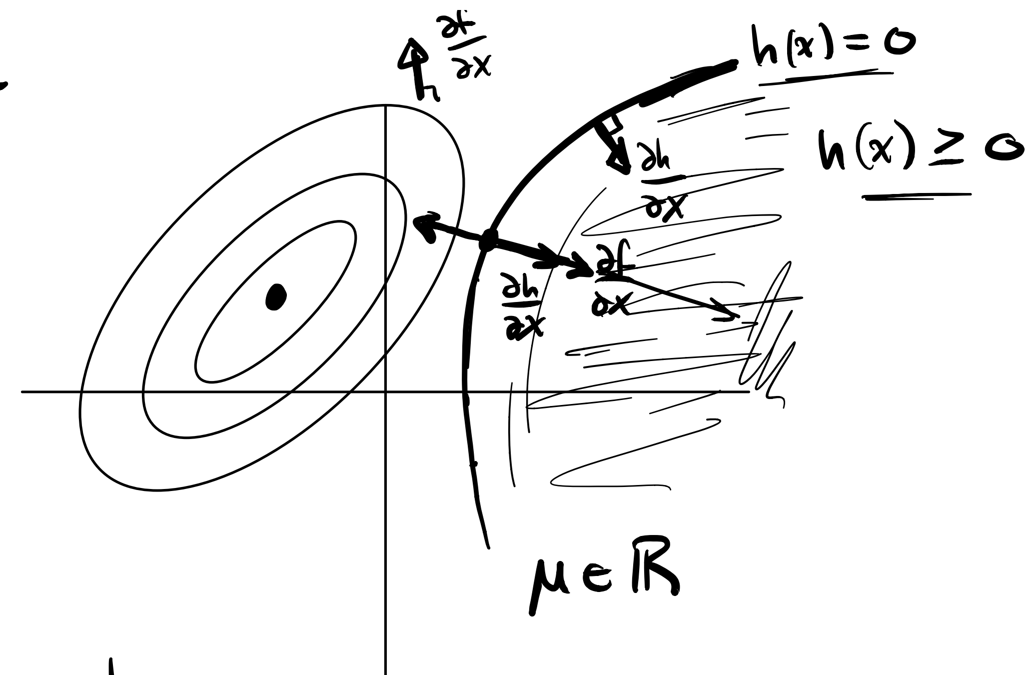
$$\frac{\partial \mathcal{L}}{\partial v} = g(x) = 0$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = h(x) \geq 0$$

$\rightarrow$ feasibility

$\underbrace{\mu \geq 0}_{\text{feasibility}} \qquad \underbrace{\mu^T h(x) = 0}_{\text{complementary slackness}}$

Picture



$\frac{\partial f}{\partial x} = \mu \frac{\partial h}{\partial x}$ $\leftarrow$ $\mu \geq 0$

if $h(x) > 0 \Rightarrow \mu = 0$
if $h(x) = 0 \Rightarrow \mu \geq 0$ } math switch

$\Updownarrow$

$\mu \geq 0, \; h(x) \geq 0, \; \mu h(x) = 0$

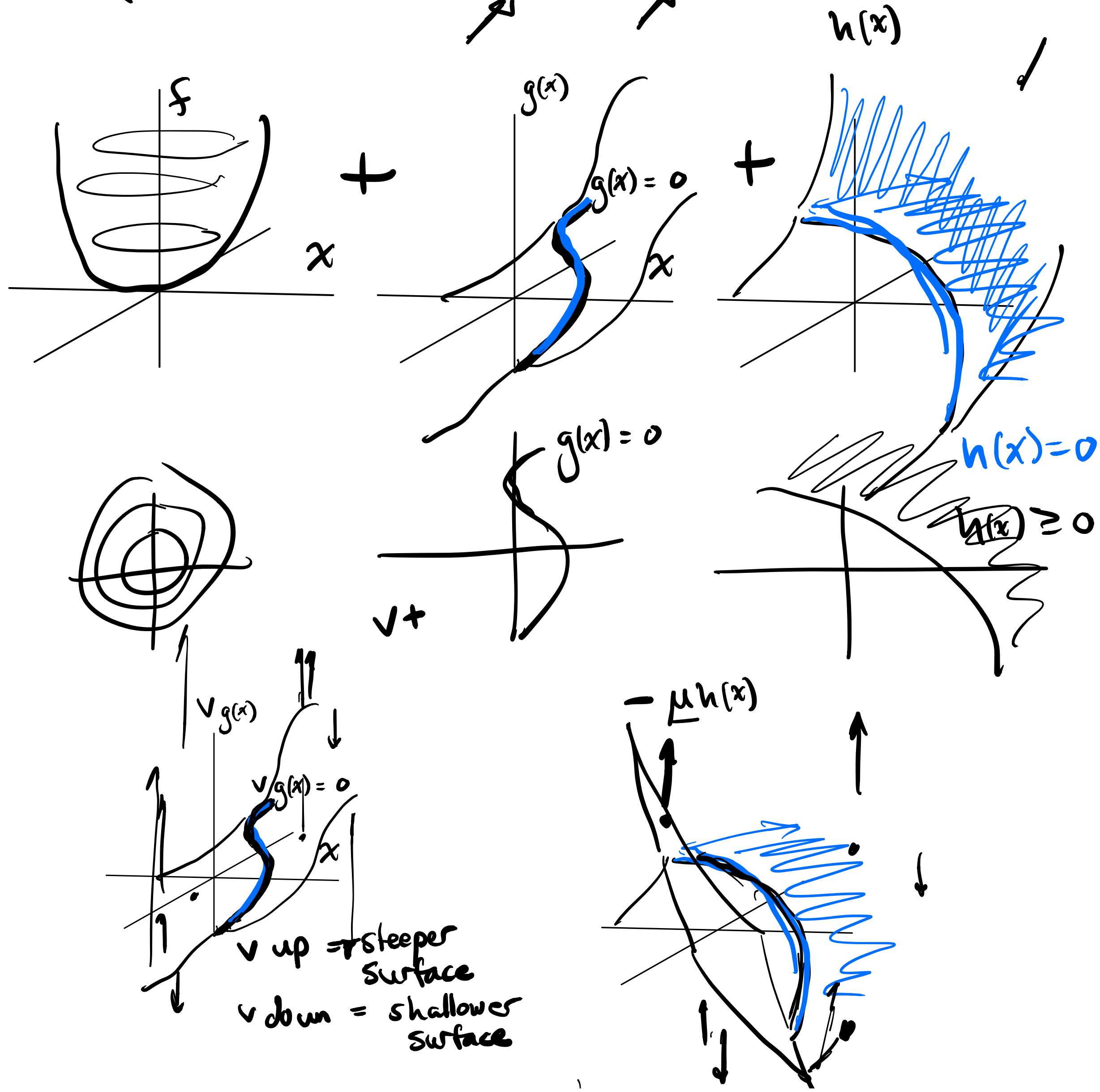
$x, s : \Rightarrow$ primal variables

$v, \mu, \tau : \Rightarrow$ dual variables

Game Theory Interpretation of Lagrangian

$$\min_x f(x) \quad \text{s.t.} \quad g(x)=0,\ h(x)\geq 0 \;=\; \min_x \max_{v,\mu\geq 0} \mathcal{L}(x,v,\mu)$$

$$\mathcal{L}(x,v,\mu) = f(x) + v^T g(x) - \mu^T h(x)$$



$g(x)=0$

$$\min_{x} \max_{v,\mu \geq 0} \; \underbrace{f(x)}_{\text{obj. surf.}} + \underbrace{v^T g(x)}_{\text{tilted } g \text{ surface}} - \underbrace{\mu^T h(x)}_{\text{tilted } h \text{ surface}}$$

$\underline{\mu}$

x: location

$v, \mu \geq 0$: tilt of surfaces

$v\,g(x)$: seesaw

$\mu\, h(x)$: one sided seesaw

$\mu \geq 0$

$h(x) \geq 0$

dual opt prob.

$g(x) = 0$

$$\boxed{\max_{v,\mu \geq 0} \min_{x} \mathcal{L}} \leq \boxed{\min_{\underline{x}} \max_{\underline{v,\mu \geq 0}} \mathcal{L}}$$

primal opt. prob

simple discrete example

$\max_{v} \min_{x} \mathcal{L} \leftarrow x$

$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\mathcal{L}=1$	$\mathcal{L}=10$
$\mathcal{L}=20$	$\mathcal{L}=2$

$\leftarrow v \rightarrow$

$\max_{v} \mathcal{L}$

$\min_{x} \max_{v} \mathcal{L}$

$[20 \;\; 10]$